特集概説

# ビッグデータを価値に変えるNECのITインフラ

近年、インターネットやICTの普及に伴い、社会全体に蓄積されるデータの総量が劇的に増大しています。一方で、大量に眠ったまま価値を発揮されないデータが大半を占めているともいわれています。
そこで本稿では「ビッグデータ」が生み出す新たな課題を整理し、それに対してNECが提供できる付加価値ビジネス、また、それを支える弊社の技術と製品群について、全体像を概観しながら説明します。

第三ITソフトウェア事業部
事業部長
伊藤 晃徳

## 1 まえがき

近年、インターネットやICTの普及に伴い、社会全体に蓄積されるデータの総量が劇的に増大しています。2020年には2006年の約200倍ものデータ量が世の中を流通するようになると予測されています[1)]。ところが、総務省の調査によると、企業内に保存されたデータの9割以上が活用されずに「宝の持ち腐れ」になっているといわれています[2)]。

企業が保有する大量データを掘り起こし、戦略立案に活用することで、新しいビジネスの機会を見出せることは、これまでも指摘されていました。これからは、更に加速するデータの流れ、増加するデータの量、多様化するデータの形式に対応していくことが必須となります。これまで活用できていなかった大量データを改めて活用できるようにし、そこから新たな知見や付加価値を導き出して新規事業に結びつけることが、今後の企業戦略における大きな課題となるでしょう。

また、近年はスマートフォンや携帯電話などモバイル端末や、ユビキタスセンサの普及による、トラフィックの急激な増加が重要な課題となっています。総務省の調査では、国内モバイルトラフィックは2010年～2011年の1年間で約2倍になっていると報告されています[3)]。これらトラフィックの増加への対応は、もはや通信キャリアのみの課題ではありません。モバイル端末やセンサを活用して情報やサービスを提供する企業すべてにかかわる課題になっているのです。

このように激増するデータは「ビッグデータ」と呼ばれ、現在注目を集めています。

本稿では、ビッグデータが生み出す新たな課題と、それに付随して生じるビジネスチャンス、及び、課題を解決するためのNECの取り組みの全体像について説明します。

## 2 ビッグデータ時代の到来と企業の課題

ビッグデータは、既にあらゆる事業領域に存在しています。マッキンゼーの調査によれば、2009年時点で保有するデータ量が多い業界のトップ3は、製造業、官公庁、通信業となっています[4)]。保有データ量が最も少ない業界でさえ、51ペタバイト（PB）ものデータが存在するといわれています。つまり、あらゆる業界において、ビッグデータへの対応が、競合との差異

化を推進し、事業の成否を分ける鍵を握っているのです。

ビッグデータの分析は、従来のビジネス・インテリジェンス（BI）とは異なり、ある業種の専門家がその業種に関するデータを分析するだけでなく、さまざまな業種・職種のユーザーが多様な観点から分析を行う必要があります。弊社では、ビッグデータ時代の到来を予測し、さまざまなデータを容易に解析できる種々の基盤技術・製品の実現に取り組んできました。次章では、ビッグデータの活用がもたらす新たなビジネスの姿について紹介します。

## 3 ビッグデータの活用で生まれるビジネスチャンス

従来のデータ活用の現場では、データの活用目的、収集方法、フォーマット、処理フローなど、さまざまな条件をあらかじめ決めてから、そのためのシステムを設計する必要がありました。しかし、ビッグデータを活用し、従来にない切り口での知見や付加価値を導き出すためには、当初想定されていなかった用途でデータを掘り出して活用したり、構造化されていない非定型データを扱ったり、活用形態にあわせてリアルタイムやバッチでの処理フローを柔軟に設定したりする必要があります。したがって、従来のデータ活用に比べてビッグデータ活用では、データを扱う環境の自由度が圧倒的に拡大します。

このように柔軟な環境でビッグデータを活用することで実現できる新たな付加価値としては、次の3つが挙げられます。

**(1) 情報提供の迅速化**

圧倒的な大量データをリアルタイムに処理するビジネスが可能になります。例えば、株式や外国為替相場などリアルタイムに変動する大量の情報を高速処理し、個人投資家や機関投資家に対し、ミリ秒単位で配信するサービスが挙げられます（**図1**）。

**(2) 新たな関係性の発見**

従来利用していなかった蓄積データから、新たな関係性・規則性を見つけ出すビジネスが実現できます。例えば、車載端末などのセンサ情報や、道路の混雑状況、工事・事故情報を収集・分析することによって、最適ルートや部品交換時期、故障診断などをドライバやディーラ、メーカに提供する、きめ細かな交通情報サービスが考えられます（**図2**）。

**(3) 将来の予見・予測**

人や物事の振る舞いを先回りして予測し、適切な対処をプロアクティブに行うビジネスが可能になります。例えば、プラント内に設置されたさまざまなセンサからの微小振動などの大量データを分析して、正常稼働時の振る舞いモデルを自動的に導出、その正常モデルとリアルタイムデータとの比較を行うことで、プラントの異常予兆を監視・検知するサービスが挙げられます（**図3**）。

これまで弊社では、ビッグデータ関連製品・技術を続々と投入してきており、大量イベント処理、画像データ認識、センサからのデータ収集など、多彩な領域で豊富な経験や実績を重ねています（**図4**）。

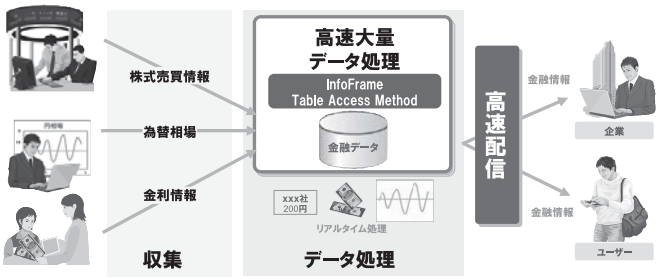


図1　金融情報の高速配信サービスの例

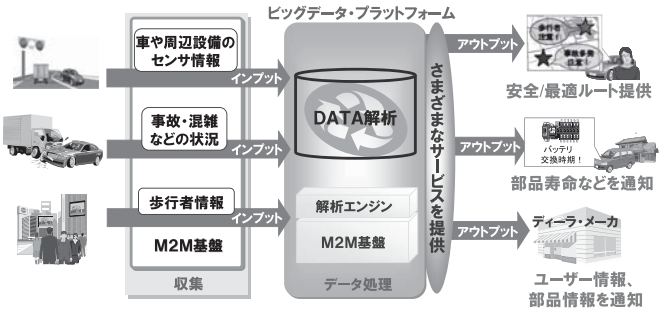


図2　センサを使った交通情報サービスの例

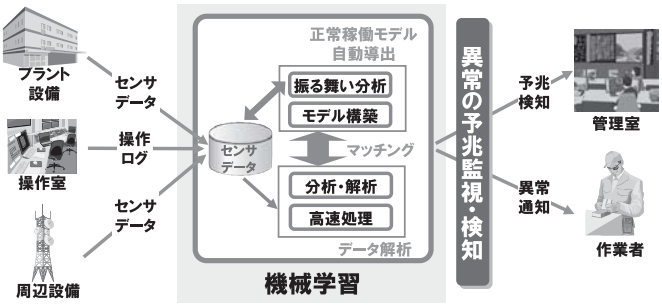


図3　プラント・建造物の異常予兆監視サービスの例

図4 NECのビッグデータにおける主な実績

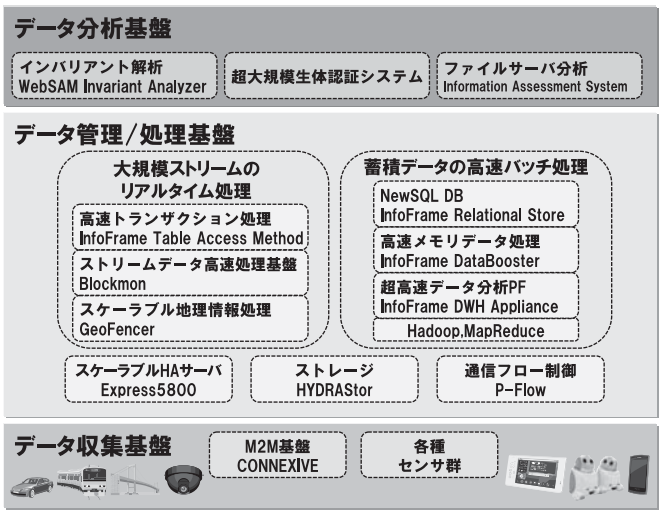


図5 NECのビッグデータデータ関連製品・技術マップ

## 4 NECのビッグデータ関連製品・技術

センサデータを含むあらゆる情報が電子化され、データ量が爆発的に増加し、かつ、さまざまなタイプのデータを収集・処理・分析する必要性が飛躍的に高まっています。ビッグデータ活用における技術的な課題は、次の3つに大別できます。

⑴ センサ・デバイスから発生する大量データの収集
⑵ 急増するデータに対応するデータ管理/処理
⑶ 画像や音声などさまざまなタイプのデータの分析

これらの課題に対応する形で、ビッグデータビジネスを支えるプラットフォームのアーキテクチャは、⑴データ収集基盤、⑵データ管理/処理基盤、⑶データ分析基盤の3層構造で構成されています（**図5**）。弊社では2006年頃よりビッグデータを強く意識した技術開発に順次取り組み始めており、それぞれの層について、既に多彩なビッグデータ処理製品を体系化しています。これらの製品群に対し、顧客のニーズに応じて最適な組み合わせで柔軟にソリューションを実現できます。ここでは、各基盤層における製品群について紹介します。

### (1) データ収集基盤

データ収集基盤層は、さまざまな種類のセンサからの入力に対し受け口を複数準備するとともに、センサデータへのアクセス方法を統一することで、データ収集を容易にします。

このレイヤにおける製品群としては、異なるアクセス方法のセンサを統一的に扱えるようにするM2Mプラットフォーム「CONNEXIVE」、農業センサ、放射能センサ、超高感度振動センサ、スマート分電盤など実世界情報をリアルタイムに取得可能な各種センサ群などが挙げられます。

「CONNEXIVE」は、あらゆるモノをつなぎ、さまざまな情報を収集・蓄積するプラットフォームです。機器ごとに異なるインタフェースの違いを吸収する汎用的なアプリケーションインタフェースを提供可能で、機器間の差異を意識せず、アプリケーションを構築することが可能です。

また、各種センサの中でも特に超高感度振動センサは、独自の振動拡大機構を採用することにより、従来の振動センサと比較して約20倍の高感度で広い周波数の振動を計測可能にし、更に約10分の1の価格という低コスト化を実現しています。

### (2) データ管理/処理基盤

データ管理/処理基盤層は用途に応じて、「大規模ストリームのリアルタイム処理」と「蓄積データの高速バッチ処理」の2つの高速大量データ処理方式を提供します。これによって、現実世界の動きに迅速に対応したリアルタイム処理と、蓄積されたデータに対する分散処理でパターン発見を行うことができます。また、データ管理/処理基盤層では、大量データ処理を行うためのハードウェア群として、スケールアウトにより大量データを効率よく処理できるサーバ、ストレージ、フロー制御も提供します。

大規模ストリームのリアルタイム処理を実現する製品群としては、「InfoFrame Table Access Method（TAM）」やストリームデータ高速処理基盤「Blockmon」が挙げられます。TAMは大量データをメモリテーブル上で連続処理可能であり、従来型基幹業務システム（ピーク時 数千件/秒）の約10倍の処理性能を実現しており、大量トランザクションのリアルタイム処理が必要とされるシステムに採用されています。また、「Blockmon」は、弊社の欧州研究所で研究開発が進められているソフトウェア基

盤であり、モジュールベースの柔軟なアーキテクチャを採用しています。そのため、トラフィックのモニタリングや通信事業者向けの不正利用ユーザー検知、企業内トラフィックの統計分析など、ストリームデータのさまざまな分析に適用できます。

蓄積データへの高速バッチ処理を実現する製品群としては、「InfoFrame Relational Store（IRS）」や「InfoFrame DataBooster」、「InfoFrame DWH Appliance」などが挙げられます。IRSは既存のリレーショナルデータベース（RDB）技術とスケールアウトの特長を持つキーバリューストア（KVS）技術のメリットを併せ持つ、ビッグデータ時代の新たなデータベースソフトウェアです。IRSの特長は次の3点です。

1) スモールスタートが可能なアーキテクチャにより初期導入コストを従来比50%削減
2) 国際標準規格の開発言語SQLに対応し、既存の設計資産の有効活用が可能
3) 独自技術「マイクロシャード」の採用により、基幹業務にも使用できる高信頼性を実現

また、「InfoFrame DataBooster」は、蓄積データの高速メモリ処理により、バッチ処理の時間を大幅に短縮したり、実行サイクルを（月次から日次へ）短縮させることが可能です。「InfoFrame DWH Appliance」は、弊社の高性能・高信頼なサーバとストレージに、Netezzaの高度なデータ検索/分析ソフトウェアを 組み合わせて最適化したデータウェアハウス・アプライアンス製品であり、従来の大規模DWHシステムと比較して約10〜100倍のパフォーマンスを実現しています。

サーバ製品「Express5800」の「スケーラブルHAサーバ」は、大容量メモリ搭載により高速メモリ処理を実現しています。「SIGMABLADE」はスケールアウト型システムにより、容易にノード追加が可能です。また、ストレージ製品である「iStorage Mシリーズ」はデータ処理の高速化のために、アクセス頻度に応じた自律データ最適配置を実現しています。更に、ネットワーク製品である「UNIVERGE PFシリーズ」では、世界で始めて次世代ネットワーク技術「OpenFlow」を採用し、データの内容に応じて最適な経路のフロー制御を実現しています。

### (3) データ分析基盤

データ分析基盤層は、画像や音声などの非構造化データに対する高い分析・解析技術やノウハウを提供します。分析・解析技術の例としては、指紋照合や顔照合などの生体認証、機械学習を用いた故障検知などが挙げられます。

弊社の指紋照合エンジンは、他人許容率0.00001%以下の世界トップの性能を実現しており、世界最大級（4,500万人規模）の南アフリカ共和国の国民IDシステムHANISに採用されています。

また、顔検出/顔照合ソフトウェア「NeoFace」も学習型超解像技術と非正面顔認証技術により、他社と比較して10倍の高精度という世界トップの性能を実現しています。

更に、機械学習を用いた故障検知技術は、大量の観測データからシステム全体の不変性（インバリアンツ）モデルを自動生成し、このモデルとリアルタイムの観測データを比較することでシステムの異常を検知することができます。この技術は運用管理製品「WebSAM Invariant Analyzer」に組み込まれています。

## 5 おわりに

以上、ビッグデータ時代の課題、及びビジネスチャンスと、それに対する弊社の取り組みとして、データ収集基盤、データ管理/処理基盤、データ分析基盤の全体アーキテクチャの概要を説明してきました。

本特集では、弊社のビッグデータソリューションを構成する各種製品群や、それを支える研究所の先進技術などについて、詳細に述べています。弊社は今後とも、これまで培ってきた大量・非構造化データの処理技術を生かしながら、クラウドサービスや最先端ICT製品の提供などを通じ、お客様企業の競争力強化のためのソリューション、サービスを提供し続けます。

*Hadoopは、The Apache Software Foundationの登録商標または商標です。

*Netezzaは、International Business Machines Corporation の登録商標または商標です。


**参考文献**

1) 喜連川 優：“ビッグデータ時代の到来,” DIGITAL GOVERNMENT, 2012.5
http://e-public.nttdata.co.jp/topics_detail2/id=658
2) 総務省：“我が国の情報通信市場の実態と情報流通量の計量に関する調査研究結果(平成21年度) ―情報流通インデックスの計量―,” 2011.8
http://www.soumu.go.jp/menu_news/s-news/01iicp01_01000004.html
3) 総務省：“我が国の移動通信トラヒックの現状,” 情報通信統計データベース, 2012.3
http://www.soumu.go.jp/johotsusintokei/field/tsuushin06.html
4) McKinsey & Company：“Big data: The next frontier for innovation, competition, and productivity,” 2011.5
http://www.mckinsey.com/insights/mgi/research/technology_and_innovation/big_data_the_next_frontier_for_innovation

# NEC 技報のご案内

NEC技報の論文をご覧いただきありがとうございます。
ご興味がありましたら、関連する他の論文もご一読ください。

**NEC技報WEBサイトはこちら**

NEC技報(日本語)　　NEC Technical Journal(英語)

## Vol.65 No.2　ビッグデータ活用を支える基盤技術・ソリューション特集

ビッグデータ活用を支える基盤技術・ソリューション特集によせて
ビッグデータを価値に変えるNECのITインフラ

### ◇ 特集論文

**データ管理 / 処理基盤**

超高速データ分析プラットフォーム「InfoFrame DWH Appliance」
SDN技術で通信フローを制御する「UNIVERGE PFシリーズ」
大量データをリアルタイムに処理する「InfoFrame Table Access Method」
大量データを高速に処理する「InfoFrame DataBooster」
ビッグデータの活用に最適なスケールアウト型新データベース「InfoFrame Relational Store」
高い信頼性と拡張性を実現したExpress5800/スケーラブルHAサーバ
大規模データ処理に対するOSS Hadoopの活用
大容量・高信頼グリッドストレージiStorage HSシリーズ(HYDRAstor)

**データ分析基盤**

ファイルサーバのデータ整理・活用を支援する「Information Assessment System」
超大規模バイオメトリック認証システムとその実現
WebSAMの分析技術と応用例～インバリアント分析の特長と適用領域～

**データ収集基盤**

スマートな社会を実現するM2Mとビッグデータ
微小な振動を検知する超高感度振動センサ技術開発とその応用

**ビッグデータ処理を支える先進技術**

多次元範囲検索を可能とするキーバリューストア「MD-HBase」
高倍率・高精細を実現する事例ベースの学習型超解像方式
ビッグデータ活用のためのテキスト分析技術
ビッグデータ時代の最先端データマイニング
ジオタグ付きデータをクラウドでスケーラブルに処理するジオフェンシングシステム
柔軟性と高性能を備えたビッグデータ・ストリーム分析プラットフォーム「Blockmon」とその使用事例

### ◇ 普通論文

地デジTVを活用した「まちづくりコミュニティ形成支援システム」

### ◇ NEC Information

**NEWS**

スケールアウト型新データベース「InfoFrame Relational Store」が2つの賞を受賞

**Vol.65 No.2**
**(2012年9月)**

特集TOP